# Modeling Relative Sophistication of Problem-Solving Strategies in Early Mathematics: A Novel Hurdle Ordinal Logit Approach

Pavel Chernyavskiy[2]
pchern@virginia.edu

Carson L. Keeter, Traci S. Kutaka, Julie Sarama, Douglas H. Clements

*August 12, 2021*

# MOTIVATION

- Greater sophistication of problem-solving strategies linked to greater **concurrent and future achievement** (Geary et al., 2017; Siegler, 1988), deeper **conceptual understanding** (Chu et al., 2018; Siegler and Shrager, 1984)

- **Sophistication** of problem-solving strategies largely ignored in education intervention studies
  - Primary focus: **correctly** answering items on of an assessment
  - Item Response Theory models to analyze (e.g., 1-PL, 2-PL, Graded Response, etc.)

# MOTIVATION

- Coding problem-solving behavior is **complex and time-intensive**:
  - Must establish a **research-based ordinal scale** for relative sophistication (here: Length-measurement Learning Trajectories; Sarama and Clements, 2009/2014)
  - Team of 4 coders strive for high (> 90%) inter-rater reliability

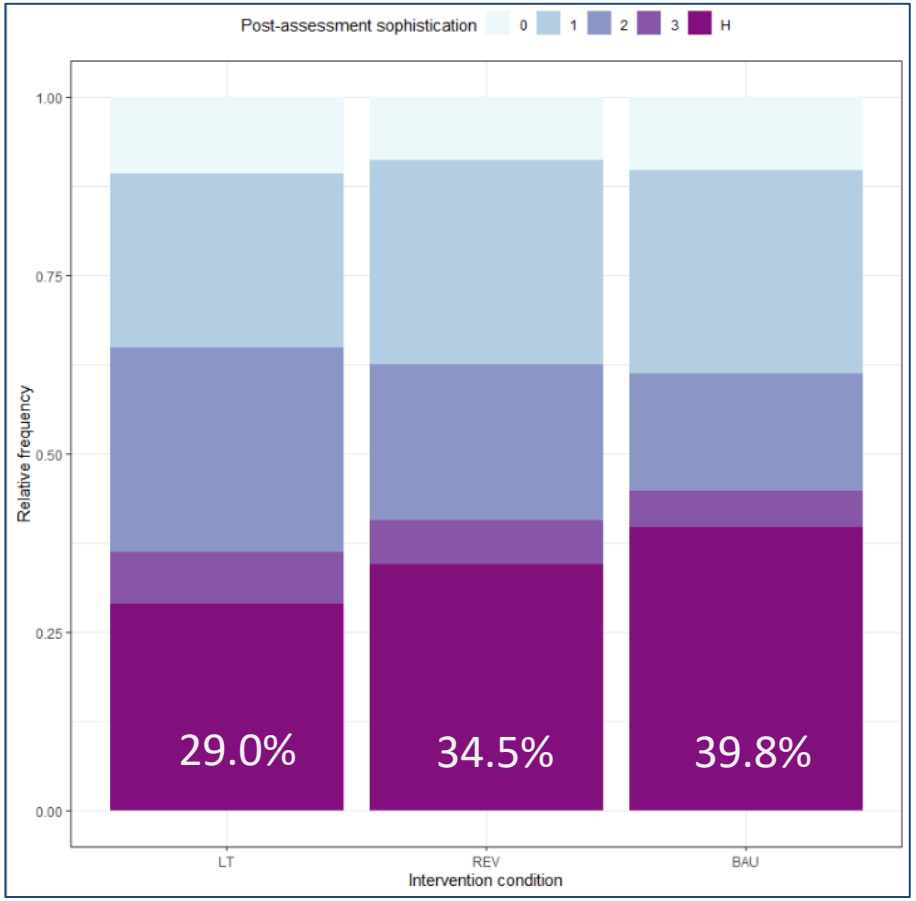
- Some behaviors **cannot** be mapped to the sophistication scale:
  - **Non-codable** and **Non-detectable** behaviors
  - Not a case of truncation or censoring

- Currently these data points are **removed** prior to analysis:
  - **Reduces power** if subjects must be excluded
  - **Detrimental to psychometric functioning** if items must be excluded

# MOTIVATION

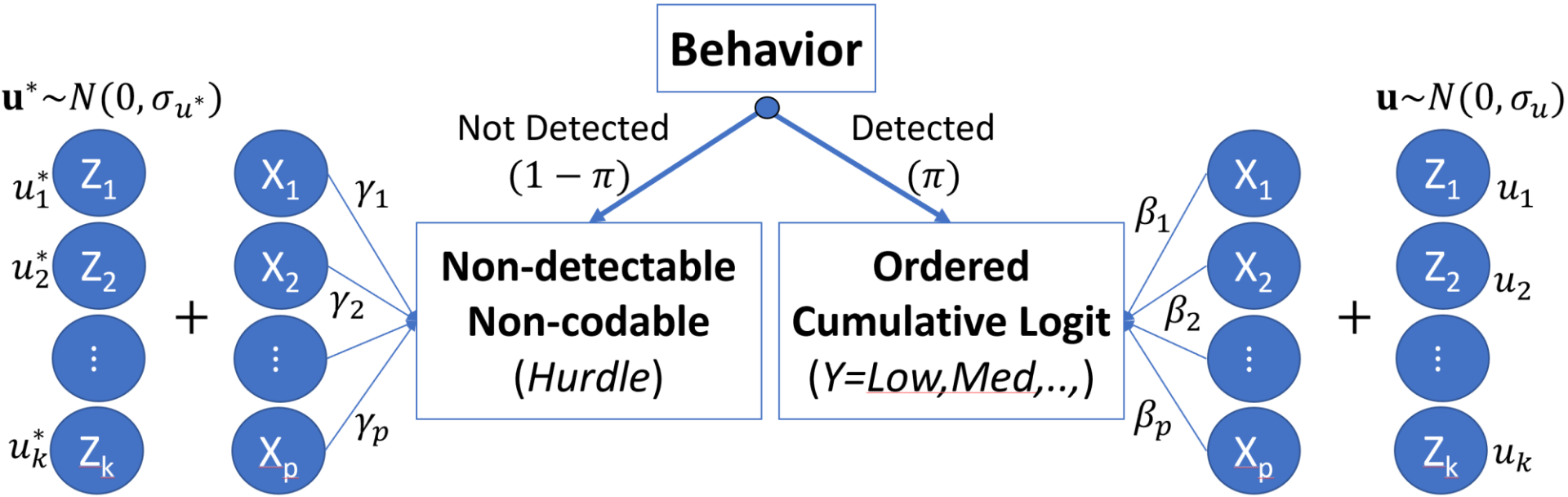
- **Motivating dataset**: early childhood intervention
  - 186 kindergarteners in a Mountain-West U.S. state
  - 26 items assessing length measurement
  - 3 experimental conditions
  - N = 186 across 16 classes in 6 schools

- **Substantial potential** for data loss:
  - >30% non-detectable, non-codable behaviors over experimental conditions
  - 7 of 26 items feature >50% non-detectable, non-codable behaviors



# HURDLE ORDINAL LOGIT

- **Proposal**: model relative sophistication **conditional** on an instance of a detected problem-solving behavior
  - **Hurdle family** of statistical models



# HURDLE ORDINAL LOGIT

- **Outcome**: $Y_{ijk}$ denotes one of *C* ordered categorical responses $c = 1, \ldots, C$ employed by *ith* subject, on the *jth* item, in *kth* classroom
  - $\pi_{ijk}$ = probability of recording a detectable, codable behavior
  - $c = H$ denotes the case where a behavior was **not** detected

- Latent variable formulation of Cumulative Logit an attractive feature:
  - **Latent variable**: latent problem-solving sophistication

$$P(Y_{ijk} = c|x_{ijk}) = \begin{cases} c = H: & 1 - \pi_{ijk}|x_{ijk} \\ c \neq H: & \pi_{ijk}|x_{ijk} \times \underbrace{\left[P(Y_{ijk} \leq c + 1|x_{ijk}) - P(Y_{ijk} \leq c|x_{ijk})\right]}_{\text{Ordered Cumulative Logit}} \end{cases}$$

# HURDLE ORDINAL LOGIT

- **Outcome**: $Y_{ijk}$ denotes one of *C* ordered categorical responses $c = 1, \ldots, C$ employed by *ith* subject, on the *jth* item, in *kth* classroom
  - $\pi_{ijk}$ = probability of recording a detectable, codable behavior
  - $c = H$ denotes the case where a behavior was **not** detected

Probability of Detection

$$\log\left(\frac{\pi_{ijk}}{1-\pi_{ijk}}\right) = \alpha + x'_{ijk}\boldsymbol{\gamma} + u_i^* + v_j^* + w_k^*$$

$$u_i^* \sim N(0, \sigma_u^*) \quad v_j^* \sim N(0, \sigma_v^*) \quad w_k^* \sim N(0, \sigma_w^*)$$

$$\{\sigma_u^*, \sigma_v^*, \sigma_w^*\} \sim N^+(0, 2.5)$$

$$\alpha \sim N(0, 2.5)$$

$$\boldsymbol{\gamma} = \{\gamma_1 \ldots \gamma_P\} \sim N(0, 2.0)$$

Sophistication | Strategy Detected

$$\log\left(\frac{\mathrm{P}(Y_{ijk} \le c)}{\mathrm{P}(Y_{ijk} > c)}\right) = \theta_c - \left(x'_{ijk}\boldsymbol{\beta} + u_i + v_j + w_k\right)$$

$$u_i \sim N(0, \sigma_u) \quad v_j \sim N(0, \sigma_v) \quad w_k \sim N(0, \sigma_w)$$

$$\{\sigma_u, \sigma_v, \sigma_w\} \sim N^+(0, 1.5)$$

$$\theta_c \sim N(0, 2.5)$$

$$\boldsymbol{\beta} = \{\beta_1 \ldots \beta_P\} \sim N(0, 1.5)$$

# ESTIMATION

- No-U-Turn Hamiltonian Monte Carlo (NUTS HMC) implemented via Stan software in R 4.0
  - 4000 samples after 1000-teration warmup across 4 MCMC chains
  - Effective Sample Size > 500 for all parameters
  - Robust MCMC sampling and convergence

- Final model selected using Leave-One-Out Information Criterion (LOOIC; Vehtari et al., 2017) and Watanabe-Akaike Information Criterion (WAIC; Watanabe, 2013)

# RESULTS

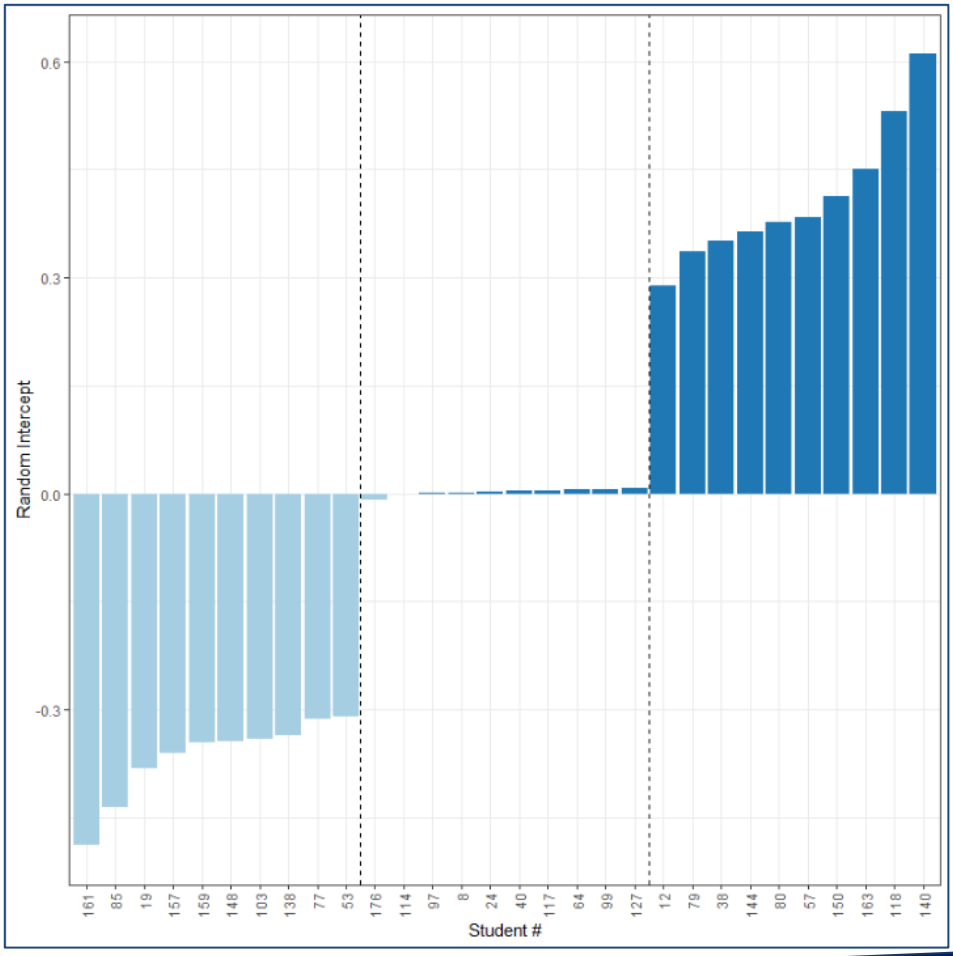
- Estimated Odds Ratios (95% Credible Intervals) :

| Covariates* | Detection | Greater Sophistication \| Detected |
|---|---|---|
| Learning Trajectory condition | **2.20 (1.62, 2.97)** | **2.44 (1.58, 4.14)** |
| Reverse condition | **1.49 (1.09, 2.08)** | **1.80 (1.28, 2.83)** |
| Business-As-Usual condition | *ref.* | *ref.* |
| Boys | 1.13 (0.86, 1.48) | 1.04 (0.81, 1.35) |
| Girls | *ref.* | *ref.* |
| Private school | 1.19 (0.69, 2.12) | **1.67 (1.16, 2.69)** |
| Public school | *ref.* | *ref.* |
| **: Adjusted for pre-sophistication Rasch score* | | |

- Hurdle component accounts for likelihood of **physical action on objects** (embodied cognition)
- Greater nuance for intervention efficacy

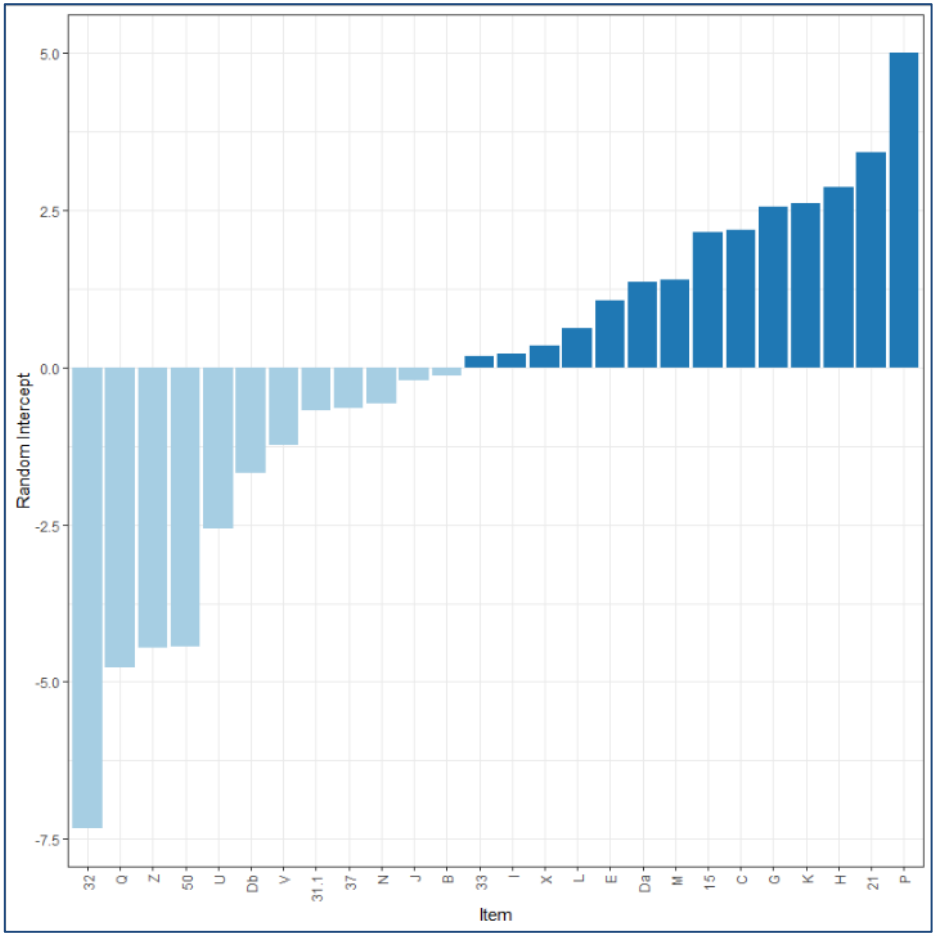
# RESULTS

- Random effects allow for scoring of **subjects** and items:
  - Student random intercepts measure **latent sophistication**, similar to the Rasch score
- Example: bottom/middle/top 10 students
- Note the relatively small magnitude



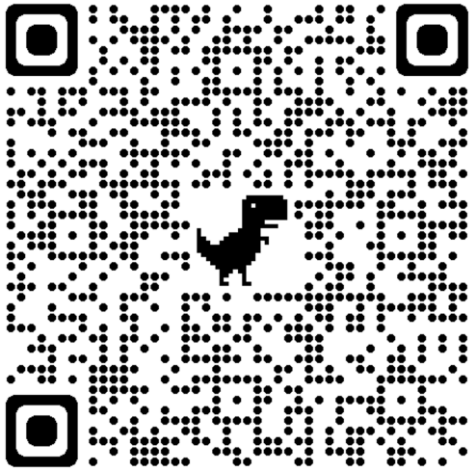
# RESULTS

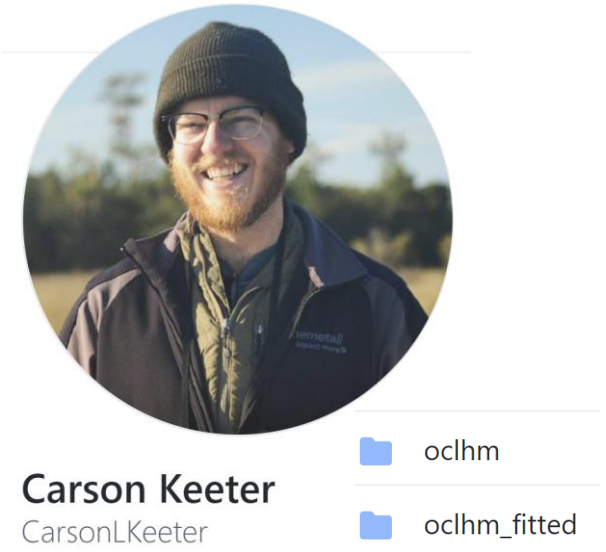
- Random effects allow for scoring of subjects and **items**:
  - Item random intercepts measure **latent difficulty**, similar to the 1-PL difficulty parameter
- Example: items ranked by estimated latent difficulty
- Note the relatively large magnitude



# CODE AND DATA AVAILABILITY

- Model estimation and predictions implemented in R functions `oclhm` and `oclhm_fitted`:
  - Available on GitHub:






- Data will be made available following an embargo
  - Toy dataset and package vignette forthcoming!


University of Virginia